



TRANSMITTAL FORM

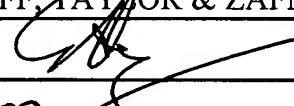
(to be used for all correspondence after initial filing)

		Application No.	10/663,476
		Filing Date	September 15, 2003
		First Named Inventor	Jong-Arm Jun
		Art Unit	2616
		Examiner Name	Xavier S. Wong
Total Number of Pages in This Submission	7	Attorney Docket Number	3364P136

ENCLOSURES (check all that apply)

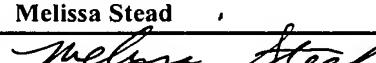
<input checked="" type="checkbox"/> Fee Transmittal Form <input type="checkbox"/> Fee Attached <input type="checkbox"/> Amendment / Reply <input type="checkbox"/> After Final <input checked="" type="checkbox"/> Affidavits/declaration(s) <input type="checkbox"/> Extension of Time Request <input type="checkbox"/> Express Abandonment Request <input type="checkbox"/> Information Disclosure Statement <input type="checkbox"/> PTO/SB/08 <input checked="" type="checkbox"/> Certified Copy of Priority Document(s) <input type="checkbox"/> Response to Missing Parts/ Incomplete Application <input type="checkbox"/> Basic Filing Fee <input type="checkbox"/> Declaration/POA <input type="checkbox"/> Response to Missing Parts under 37 CFR 1.52 or 1.53	<input type="checkbox"/> Drawing(s) <input type="checkbox"/> Licensing-related Papers <input type="checkbox"/> Petition <input type="checkbox"/> Petition to Convert a Provisional Application <input type="checkbox"/> Power of Attorney, Revocation Change of Correspondence Address <input type="checkbox"/> Terminal Disclaimer <input type="checkbox"/> Request for Refund <input type="checkbox"/> CD, Number of CD(s) <input type="checkbox"/> Landscape Table on CD	<input type="checkbox"/> After Allowance Communication to TC <input type="checkbox"/> Appeal Communication to Board of Appeals and Interferences <input type="checkbox"/> Appeal Communication to TC (Appeal Notice, Brief, Reply Brief) <input type="checkbox"/> Proprietary Information <input type="checkbox"/> Status Letter <input checked="" type="checkbox"/> Other Enclosure(s) (please identify below): <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Request for Priority; return postcard</div>	
			<input type="checkbox"/>

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm or Individual name	Eric S. Hyman, Reg. No. 30,139 BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
Signature	
Date	10/3/07

CERTIFICATE OF MAILING/TRANSMISSION

I hereby certify that this correspondence is being deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Typed or printed name	Melissa Stead
Signature	
Date	10-3-07



FEE TRANSMITTAL for FY 2006

Patent fees are subject to annual revision.

Complete if Known

Application Number	10/663,476
Filing Date	September 15, 2003
First Named Inventor	Jong-Arm Jun
Examiner Name	Xavier S. Wong
Art Unit	2616
Attorney Docket No.	3364P136

Applicant claims small entity status. See 37 CFR 1.27.

TOTAL AMOUNT OF PAYMENT (\$)

METHOD OF PAYMENT (check all that apply)

Check Credit card Money Order None Other (please identify): _____

Deposit Account Deposit Account Number: 02-2666 Deposit Account Name: Blakely, Sokoloff, Taylor & Zafman LLP

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

Charge fee(s) indicated below Charge fee(s) indicated below, except for the filing fee
 Charge any additional fee(s) or underpayment of fee(s) Credit any overpayments
 under 37 CFR §§ 1.16, 1.17, 1.18 and 1.20.

FEE CALCULATION

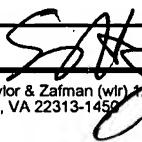
Large Entity Small Entity

Fee Code	Fee (\$)	Fee Code	Fee (\$)	Fee Description	Fee Paid
1051		2051		Surcharge - late filing fee or oath	
1052		2052		Surcharge - late provisional filing fee or cover sheet.	
2053		2053		Non-English specification	
1251	120	2251	60	Extension for reply within first month	
1252	460	2252	230	Extension for reply within second month	
1253	1,050	2253	525	Extension for reply within third month	
1254	1,640	2254	820	Extension for reply within fourth month	
1255	2,230	2255	1,115	Extension for reply within fifth month	
1401	510	2401	255	Notice of Appeal	
1402	510	2402	255	Filing a brief in support of an appeal	
1403	1,030	2403	515	Request for oral hearing	
1451	1,510	2451	1,510	Petition to institute a public use proceeding	
1460	130	2460	130	Petitions to the Commissioner	
1807	50	1807	50	Processing fee under 37 CFR 1.17(q)	
1806	180	1806	180	Submission of Information Disclosure Stmt	
1809	810	1809	405	Filing a submission after final rejection (37 CFR § 1.129(a))	
1810	810	2810	405	For each additional invention to be examined (37 CFR § 1.129(b))	

Other fee (specify) _____

SUBTOTAL (2) (\$)

Complete (if applicable)

Name (Print/Type)	Eric S. Hyman	Registration No. (Attorney/Agent)	30,139	Telephone	(310) 207-3800
Signature				Date	10/3/07



DOCKET NO.: 3364P136

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

JONG-ARM JUN, ET AL.

Application No.: 10/663,476

Filed: September 15, 2003

For: **SCALABLE CROSSBAR MATRIX
SWITCH AND ARBITRATION
METHOD THEREOF**

Art Group: 2616

Examiner: Xavier S. Wong

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR PRIORITY

Applicant respectfully requests a convention priority for the above-captioned application, namely:

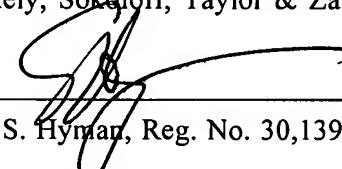
COUNTRY	APPLICATION NUMBER	DATE OF FILING
Korea	2002-0083720	24 December 2002
Korea	2003-0013079	3 March 2003

A certified copy of the document is being submitted herewith.

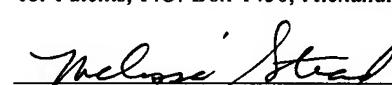
Dated: 10/3/07

1279 Oakmead Parkway
Sunnyvale, CA 94085-4040
Telephone: (310) 207-3800

Respectfully submitted,
Blakely, Sokoloff, Taylor & Zafman LLP

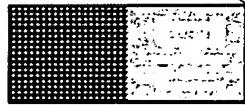

Eric S. Hyman, Reg. No. 30,139

I hereby certify that this correspondence is being deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.


Melissa Stead

10-3-07

Date



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출 원 번 호 : 10-2002-0083720

Application Number

출 원 년 월 일 : 2002년 12월 24일

Filing Date DEC 24, 2002

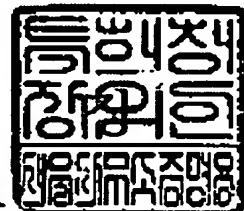
출 원 인 : 한국전자통신연구원

Applicant(s) Electronics and Telecommunications Research Institute



특 허 청

COMMISSIONER



2007년 09월 10일

◆ This certificate was issued by Korean Intellectual Property Office. Please confirm any forgery or alteration of the contents by an issue number or a barcode of the document below through the KIPOnet- Online Issue of the Certificates' menu of Korean Intellectual Property Office homepage (www.kipo.go.kr). But please notice that the confirmation by the issue number is available only for 90 days.

【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0006
【제출일자】	2002.12.24
【발명의 국문명칭】	이차원-확장형 크로스바 매트릭스 스위치용 장치
【발명의 영문명칭】	Method and Apparatus for Two-Dimensional-Scalable Crossbar Matrix Switch
【출원인】	
【명칭】	한국전자통신연구원
【출원인코드】	3-1998-007763-8
【대리인】	
【명칭】	유미특허법인
【대리인코드】	9-2001-100003-6
【지정된변리사】	이원일
【포괄위임등록번호】	2001-038431-4
【발명자】	
【성명의 국문표기】	전종암
【성명의 영문표기】	JUN, JONG ARM
【주민등록번호】	641123-1XXXXXXX
【우편번호】	305-755
【주소】	대전광역시 유성구 어은동 99번지 한빛아파트 125동 605호
【국적】	KR
【발명자】	
【성명의 국문표기】	변성혁
【성명의 영문표기】	BYUN, SUNG HYUK
【주민등록번호】	700402-1XXXXXXX

【우편번호】	305-350
【주소】	대전광역시 유성구 가정동 236-1번지
【국적】	KR
【발명자】	
【성명의 국문표기】	안병준
【성명의 영문표기】	AWN, BYUNG JUN
【주민등록번호】	620124-1XXXXXX
【우편번호】	305-755
【주소】	대전광역시 유성구 어은동 한빛아파트 129동 402호
【국적】	KR
【발명자】	
【성명의 국문표기】	남승엽
【성명의 영문표기】	NAM, SEUNG YEOB
【주민등록번호】	750821-1XXXXXX
【우편번호】	305-338
【주소】	대전광역시 유성구 구성동 373-1번지
【국적】	KR
【발명자】	
【성명의 국문표기】	성단근
【성명의 영문표기】	SUNG, DAN KEUN
【주민등록번호】	520719-1XXXXXX
【우편번호】	305-805
【주소】	대전광역시 유성구 신성동 160-1번지 한울아파트 103동 1503호
【국적】	KR
【공지예외적용대상증명서류의 내용】	
【공개형태】	논문발표
【공개일자】	2002.09.17

【취지】 특허법 제42조의 규정에 의하여 위와 같이 출원합니다.

대리인

유미특허법인 (인)

【수수료】

【기본출원료】	20	면	29,000	원
【가산출원료】	12	면	12,000	원
【우선권주장료】	0	건	0	원
【심사청구료】	0	항	0	원
【합계】			41,000	원
【감면사유】			정부출연연구기관	
【감면후 수수료】			20,500	원

【기술이전】

【기술양도】	희망
【실시권허여】	희망
【기술지도】	희망

【첨부서류】 1. 요약서 · 명세서(도면)_1통 2. 공지예외적용대상(신규성상
실의예외, 출원시의특례)규정을 적용받 기 위한 증명서류_
1통

【요약서】

【요약】

본 발명은 수백 기가급 고속 스위치에 많이 적용되고 있는 Input-Queued Crossbar Switch를 테라급 이상의 고속 대용량 스위치에 적용하기 위한 이차원-확장형 크로스바 매트릭스 스위치에 대한 구성방식 및 장치에 관한 것이다. 본 발명은 N개의 input ports와 N개의 output ports와 input port와 output port 사이의 셀을 전달하기 위한 $N \times N$ switch fabric으로 구성된다. 각 input port는 Head-Of-Line (HOL) blocking 문제의 해결을 위하여 N개의 VOQ(Virtual Output Queue)를 가지며, N 개의 VOQ는 순차적으로 n개의 VOQ 단위로 묶어서 물리적으로 분리된 L개의 VOQ 그룹을 구성한다. 여기서 n은 N의 약수이며 $L=N / n$ 이다. L개의 VOQ 그룹은 각각 독립된 switch fabric interface port를 통하여 물리적으로 분리된 L개의 XSU에 각각 연결된다. 따라서 각 input port 는 a cell time slot 동안 최대 L개의 셀을 switch fabric에 전달 가능하다. $N \times N$ 크기의 switch fabric은 L2 개의 XSU(Crossbar Switch Unit)로 구성되며, 각 XSU는 서로 독립된 arbitration을 수행 한다. 각 XSU는 n개의 input ports와 n개의 output ports를 가지며, 각 output port에는 일정크기의 셀을 저장할 수 있는 buffer가 있다. 단일 XSU는 output port 별 buffer가 셀을 받아 들일 수가 있을 경우 output port 별 독립된 arbitration을 통하여 셀을 전달한다. 각 XSU는 n 개의 독립된 input port와 n 개의 독립된 output port에 접속이 된다. N개의 Output ports는 서로 독립된 arbitration을 수행한다. 각 output port는 L개의 switch fabric interface port를

1020020083720

통하여 물리적으로 분리된 L개의 XSU에 연결되어있으며, output port arbiter는 a cell time slot 동안 L개의 XSU로부터 최대 1개의 셀을 선택하는 기능을 수행한다.

【대표도】

도 2

【색인어】

Crossbar Switch, Arbiter, Scheduler, Scalability

【명세서】

【발명의 명칭】

이차원-확장형 크로스바 매트릭스 스위치용 장치 {Method and Apparatus for Two-Dimensional-Scalable Crossbar Matrix Switch}

【도면의 간단한 설명】

- <1> 도 1은 본 발명의 $N \times N$ 크로스바 매트릭스 스위치의 구조도.
- <2> 도 2는 본 발명의 L2개의 크로스바 단위 스위치를 상호 연결하여 $N \times N$ 크로스바 매트릭스 스위치를 구성하는 연결 방식도.
- <3> 도 3은 본 발명의 XSU Arbiter에 적용 가능한 arbiter의 1단계 동작 순서도.
- <4> 도 4는 본 발명의 XSU Arbiter에 적용 가능한 arbiter의 2 단계 동작 순서도.
- <5> 도 5는 본 발명의 XSU Arbiter에 적용 가능한 arbiter의 3단계 동작 순서도.
- <6> 도 6은 본 발명의 XSU Arbiter에 적용 가능한 Grant Arbiter의 장치구현 실시예를 도시한 도면.
- <7> 도 7은 본 발명의 XSU Arbiter에 적용 가능한 Accept Arbiter의 장치구현 실시예를 도시한 도면.
- <8> 도 8은 본 발명에 적용 가능한 XSU Arbiter의 장치구현 실시예를 도시한 도면.
- <9> 도 9는 본 발명의 output port에 적용 가능한 Output Arbiter의 장치구현

실시예를 도시한 도면.

<10> 도 10은 본 발명에 적용 가능한 L2 개의 XSU Arbiter와 N 개의 Output Arbiter의 interconnection을 통한 N x N 크로스바 매트릭스 스위치 Arbiter의 장치구현 실시예를 도시한 도면.

<11> <도면의 주요 부분에 대한 부호의 설명>

<12> 1 VOQ : Virtual Output Queue

<13> 2 XSU : Crossbar Switch Unit

<14> 3 XSM : Crossbar Switch Module

<15> 4 SDB : Scheduling Decomposition Buffer

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

<16> 본 발명은 고속 대용량의 스위치를 구성하는 방식과 본 스위치를 통하여 셀을 전달하기 위한 scheduling 방식에 대한 방식이다. 본 발명은 LAN 스위치, ATM 교환기, IP router 등의 고속 대용량 크로스바 스위치로서 활용 가능하다. 좀 더 구체적으로 설명하면, 본 발명은 n이 N의 약수이고 $L = N / n$ 일 때 L2 개의 n x n Crossbar Switch Unit을 상호 연결하여 N x N 스위치를 구성하는 방식과 본 스위치의 input port에서 output port로 셀을 전달하기 위한 scheduling 방식에 대한 것이다.

<17> N x N Output-Queued Switch의 경우 특정 output port로 a time-slot에 최대 N개의 셀을 전달 할 수 있어야 하므로 스위치 fabric과 메모리의 동작 속도는 input line rate보다 N배의 속도로 동작해야 한다. Output-Queued Switch가 100%의 throughput과 QoS 지원 측면에서 우수한 특성을 가지지만 고속 스위치에 적용되지 않는 주된 이유는 이와 같은 speed-up 문제 때문이다.

<18> Input-Queued switch의 경우 스위치 fabric과 메모리의 동작 속도가 input line rate와 동일하게 동작할 수 있다는 장점 때문에 대부분의 고속 스위치에 많이 적용되고 있다. 고속 Input-Queued switch의 구조는 Input port의 경우 Head-Of-Line blocking 문제를 해결하기 위한 VOQ 형태를 가지며, 스위치 fabric의 경우 non-blocking 고속 스위치 fabric 소자인 crossbar 형태를 가지는 구조가 대부분이다. Input-Queued switch에서 input port와 output port에서의 contention 문제를 해결하여 100% 까지 throughput을 제공할 수 있는 다양한 arbitration 방식이 최근 많이 제안되고 있다. 그 중에는 parallel iterative matching (PIM) (U.S.Pat.No.5,267,235), 2-Dimensional Round Robin Matching (2DRR) (U.S.Pat.No.5,299,190), iSLIP (U.S.Pat.No.5,500,858), Dual Round Robin Matching(H. J. Chao and J-S Park, "Centralized Contention Resolution Schemes for a Large-Capacity Optical ATM switch, Proc. IEEE ATM Workshop97, Fairfax, VA, pp.10-11, May 1988) 등과 같은 알고리즘이 있다. 이와 같은 알고리즘들은 100% throughput 을 제공 가능하지만 입출력 포트의 수가 커질수록 arbitration time limitation을 가지게 되므로 대용량 고속 스위치에는 적용이 어려운 단점을

가지고 있다. 예를 들어 상용화 제품에 가장 많이 적용이 되었던 iSLIP의 경우에도 32포트 이하의 스위치에만 주로 적용이 되었다. 2.5 Gbps 포트속도를 가지는 테라급 스위치의 경우 최소 512 x 512 크기의 crossbar switch fabric이 필요하므로 이를 단일 crossbar switch fabric과 기존의 arbitration 방식을 사용하여 구성하는 것은 근본적인 문제점을 가지고 있다.

<19> 대용량 스위치 구조로서 제안되고 있는 확장 가능형 스위치 구조로는 Distribution Network기반의 channel group switching 방식(U.S.Pat.No.5,724,351)과 최근에 많이 적용되고 있는 clos 네트워크 형태의 crossbar switch 구조(F. M. Chiussi, J. G. Kneuer, and V. P. Kumar, " Low-cost Scalable Switching Solutions For Broadband Networking: The Atlanta Architecture And Chipset ", IEEE Commun. Mag., pp.4453, Dec. 1997.)가 있다. Channel group switching 방식의 경우 소용량 group 스위치 소자를 이용하여 대용량 스위치를 구성할 수 있는 장점은 있지만 입출력 포트가 많아지면 Distribution Network의 설계가 매우 복잡하여 대용량 스위치의 구성이 어려운 단점이 있다. Clos 네트워크 형태의 crossbar switch 구조의 경우 입출력 포트간 다수의 경로가 존재하므로 경로별 트래픽의 load를 분산시키기 위한 복잡한 scheduling 방식이 필요할 뿐만 아니라, 고속 입출력 포트의 수가 많아지면 이에 비례하여 고속의 crossbar switch 소자가 필요하므로 대용량 스위치의 구성에 어려운 단점이 있다.

【발명이 이루고자 하는 기술적 과제】

<20> 본 발명은 입출력 포트간 단일 경로가 존재하면서 고속의 대용량 스위치에 적용 가능한 확장가능 스위치 구조와 이러한 스위치구조에 적합한 분산형 arbitration 방식을 제안한다. 단일 경로 확장가능 스위치 구조의 경우 소규모 Crossbar Switch Unit들을 상호 연결하여 대규모 스위치를 구성하는 방식을 제안하며, 분산형 arbitration방식의 경우 충분한 arbitration time margin을 가지면서 100% 까지 throughput 을 지원할 수 있는 방식을 제안한다.

【발명의 구성】

<21> [도 1]은 본 발명의 $N \times N$ 크로스바 매트릭스 스위치의 구조를 도시한 것이다. 본 발명은 N개의 Input Port(10)와 N개의 Output Port(16)와 input port와 output port 사이의 셀을 전달하기위한 $N \times N$ switch fabric(14)으로 구성된다. 각 Input Port(10)는 Head-Of-Line (HOL) blocking문제의 해결을 위하여 N개의 Virtual Output Queue (이후 VOQ로 표기)를 가지며, N 개의 VOQ는 순차적으로 n개의 VOQ단위로 묶어서 물리적으로 분리된 L개의 VOQ 그룹(11)을 구성한다. 여기서 n은 N의 약수이며 $L=N / n$ 이다. L개의 VOQ 그룹(11)은 각각 독립된 switch fabric interface port(12)를 통하여 물리적으로 분리된 L개의 XSU(13)에 각각 연결된다. 따라서 각 input port (10)는 1 cell time slot 동안 최대 L개의 셀을 switch fabric에 전달 가능하다.

<22> $N \times N$ 크기의 switch fabric(14)은 L2 개의 crossbar switching unit (이후

XSU로 표기) (13)으로 구성되며, 각 XSU는 n 개의 독립된 input port[10]와 n 개의 독립된 Output Port(16)에 접속이 된다. 각 XSU(13)는 서로 독립된 arbitration 기능을 통하여 n개의 input ports[10]로부터 들어오는 셀을 n개의 Output Port(16)로 적절하게 전달하는 기능을 수행한다.

<23> 각 Output Port는 L개의 switch fabric interface port(15)를 통하여 물리적으로 분리된 L개의 XSU(13)에 연결되어있으며, Output Port(16)는 서로 독립된 arbitration을 수행하면서 a cell time slot 동안 L개의 XSU(13)로부터 최대 1개의 셀을 선택하여 output link로 전달하는 기능을 수행한다.

<24> [도 2]는 L2개의 크로스바 단위 스위치를 상호 연결하여 본 발명의 N x N 크로스바 매트릭스 스위치를 구성하는 내부 연결 방식을 나타낸 것이다. XSU(13)는 n 개의 input ports와 n개의 output ports를 가지는 n x n crossbar switch fabric(22)과, 각 output port에 일정크기의 셀을 저장할 수 있는 small size scheduling decomposition buffer(이후 SDB로 표기) (24)로 구성되어 있다. XSU(13)의 각 output port는 SDB (24)가 셀을 받아 들일 수가 있을 경우 XSU(13) output port별 독립된 arbitration을 통하여 해당 SDB(24)에 셀을 저장하는 기능을 제공한다. XSU(13)의 input port와 output port에 적용되는 arbitration 방식은 다양한 방식을 적용할 수 있지만 고속으로 구현하기에 가장 적합한 방식은 round-robin기반의 기반의 방식이다.

<25> 하나의 crossbar switch module (이후 XSM으로 표기) (20)은 L개의 XSU(13)으로 구성된다. 하나의 XSM (20)은 n개의 Input Port(10)로부터 셀을 받아서 N개의

Output Port(16)로 전달하는 기능을 제공한다. m 번째XSM에 연결되는 input port는 $(m \times n)$ 번째 input port로부터 $\{m \times (n+1) 1\}$ 번째 input port까지 이다. 단일 XSM과 n 개의 input ports는 물리적으로 분리된 Subsystem으로 구성할 수 있고, 이러한 L 개의 Subsystem을 이용하여 전체 $N \times N$ 시스템을 구성할 수 있는 modular한 특성을 가지고 있으므로 일반적으로 Multi-Rack으로 구성할 수 밖에 없는 대용량 스위치를 modular하게 구성할 수 있는 장점을 가지고 있다.

<26> 하나의 output port(16)은 memory-less 형태로 구성되며, L 개의 SDB(24)로부터 a time slot동안 최대 1개의 셀을 Output Arbiter(26)가 선택하여 output port로 전달하는 기능을 수행한다. j 번째 출력포트에 해당되는 SDB는 $\sum_{m=0}^{L-1} SDB(m, \left\lfloor \frac{j}{n} \right\rfloor, j \bmod n)$ 과 같이 표현된다. Output Arbiter(26)의 경우 다양한 방식을 적용할 수 있지만 고속으로 구현하기에 가장 적합한 방식은 round-robin기반의 방식이다.

<27> 본 발명의 대상인 $N \times N$ 크로스바 매트릭스 스위치의 입출력 사이의 arbitration은 two stage distributed arbitration 구조로 이루어져 있다. 첫번째 stage arbitration은 분산된 각 XSU(13)에 의해 수행이 되며, 두번째 stage arbitration은 분산된 각 Output Port(16)별로 수행이 된다. 각 stage는 internal speed-up이 없이 상호 독립적면서 concurrent한 arbitration 기능을 수행한다.

<28> 첫번째 stage arbitration은 iSLIP을 확장한 방식으로서 XSU(13)의 각 output port는 SDB(24)의 상태 정보인 credit 정보(1 : SDB가 셀을 받아들일 수 없

을 경우, 0 : SDB가 셀을 받아들일 수 있을 경우)를 사용하여 arbitration 기능을 수행한다. XSU(13)의 arbitration의 동작은 다음과 같은 3 step으로 이루어져 있다.

<29> 1)Step 1. Request. In each time slot, non-empty VOQs send a request to every output port arbiter.

<30> 2)Step 2. Grant. Each output port arbiter chooses one request in a round-robin fashion starting from the highest priority elements if its credit is 0. It then sends the grant to the selected VOQ.

<31> 3)Step 3. Accept. If an input receives grants, it accepts the one in a round-robin fashion starting from the highest priority elements.

<32> 첫번째 stage arbitration에 대해서는 [도 3], [도 4], [도 5]에서 3 step에 대한 세부 알고리즘을 설명하며, [도 6], [도 7], [도 8]에서 구체적인 구현 예를 설명한다.

<33> 두번째 stage arbitration은 분산된 각 Output Port(16)별로 수행이 되며 기본적인 알고리즘은 round-robin 방식으로서 다음과 같은 2 step으로 이루어져 있다.

<34> 1)Step 1. Request. In each time slot, non-empty SDBs send a request to every output port arbiter.

<35> 2)Step 2. Accept. Each output port arbiter chooses one request in a

round-robin fashion starting from the highest priority elements.

<36> 두번째 stage arbitration에 대해서는 [도 9]에서 구체적인 구현 예를 설명 한다.

<37> [도 3]은 본 발명의 XSU(13)에 적용 가능한 arbiter의 1단계 동작으로서 Request 신호의 생성과 관련된 예를 나타낸 것이다. 본 예는 4개의 input ports와 4개의 output ports를 가지는 4×4 XSU의 예이다. 각 input port별 Accept Arbiter(30-1) ~ (30-4)가 있으며 각 arbiter에는 port별 highest priority value를 저장하기 위한 accept register인 $a0 \sim a3$ 이 있으며 각 register의 현재 값은 0으로 가정한다. 각 output port별 Grant Arbiter (32-1) ~ (32-4)가 있으며 각 arbiter에는 port별 highest priority value를 저장하기 위한 grant register인 $g0 \sim g3$ 이 있으며 각 register의 현재 값은 0으로 가정한다. 각 input port별 non-empty VOQ의 상태에 따라서 request 신호를 생성하여 해당하는 output port에 전달 한다. 예를 들면 Input port 0의 경우 output port 0와 1에 해당하는 VOQ에 최소한 1셀 이상이 존재하여 해당 output port로 request 신호를 보낸다. XSU(13)의 각 output port는 SDB(24)의 상태 정보인 credit 정보를 가지게 되는데 본 예에서 output port 0의 경우 credit 값이 1이므로 SDB가 full인 상태를 의미하며, 이를 제외한 나머지 output port의 credit 값은 0으로서 최소한 1개 이상의 셀을 받아 들일 수가 있는 상태임을 의미한다.

<38> [도 4]는 본 발명의 XSU(13)에 적용 가능한 arbiter의 2단계 동작으로서 Grant 신호의 생성과 관련된 예를 나타낸 것이다. Output port 0의 경우 credit 값

이 1이므로 더 이상의 셀을 받아들일 수가 없으므로 grant 신호를 input port에 전달하지 않는다. 그 이외의 output port의 경우 credit 값이 0이므로 Grant Arbiter (32-1)~(32-4)에 저장되어있는 highest priority value로부터 round-robin 형태로 request신호를 검색하여 그 중 하나를 선택한다. 예를 들면 output port 3의 Grant Arbiter (32-4)의 경우 현재 highest priority value가 0이고 input port 2와 input port 3으로부터 request 신호를 받았으므로 input port 2에게 grant 신호를 보낸다. Grant Arbiter (32-1)~(32-4)의 highest priority value update는 3단계 accept 단계에서 accept를 받은 이후에 이루어지므로 본 2단계에서 Grant Arbiter (32-1) ~ (32-4)별 grant register $g0 \sim g3$ 의 값은 1단계와 동일한 상태를 유지한다.

<39> [도 5]는 본 발명의 XSU(13)에 적용 가능한 arbiter의 3단계 동작으로서 Accept 신호의 생성과 관련된 예를 나타낸 것이다. Output Port Arbiter(32-1)~(32-4)로부터 grant 신호를 받은 input port는 Accept Arbiter (30-1)~(30-4)에 저장되어있는 highest priority value로부터 round-robin 형태로 grant 신호를 검색하여 그 중 하나를 선택한다.

<40> Input port 0의 경우 output port 1에서만 grant 신호를 받았으므로 해당 포트로 accept 신호를 전달하며, 그 후 input port 0의 Accept Arbiter(30-1)의 경우 highest priority value를 2로 set하며, output port 1 의 Grant Arbiter (32-2)의 경우 highest priority value를 1로 set한다. 즉 다음 time slot에서 input port 0의 경우 output port 2에 우선권을 부여하며, output port 1의 경우 input port 1

에 우선권을 부여한다.

<41> Input port 2의 경우 현재 highest priority value가 0이고 output port 2와 output port 3으로부터 grant 신호를 받았으므로 output port 2에게 accept 신호를 보내며, 그 후 input port 2의 Accept Arbiter(30-3)의 경우 highest priority value를 3로 set하며, output port 2의 Grant Arbiter (32-3)의 경우 highest priority value를 3으로 set한다.

<42> [도 4]에서 output port 3의 Grant Arbiter (32-4)의 경우 grant 신호를 input port 2로 전달하였지만 3단계에서 input port 2로부터 accept 신호를 받지 못하였으므로 Grant Arbiter (32-4)의 highest priority value는 0으로 그대로 유지된다.

<43> [도 6]은 본 발명의 XSU(13)의 Grant Arbiter(60)에 대한 장치구현 실시 예를 나타낸 것이다. Request Signal Vector(61)가 XSU(13)의 Grant Arbiter(60)에 전달된다. Request Signal Vector(61)는 n bit 크기로 XSU(13)의 input port에서 특정output port로 전달하는 request 신호를 의미한다.

<44> Request Signal Vector(61)는 Round Robin Selector(62)에 연결된다. Round Robin Selector(62)는 Grant Arbiter g_j (65)에 저장되어 있는 highest priority value로부터 round-robin 형태로 Request Signal Vector(61)를 검색하여 처음으로 request된 신호 하나를 선택한다. Round Robin Selector(62)는 선택된 request 신호에 대한 결과로서 $\log(n)$ 크기의 신호를 출력한다. $\log(n)$ 크기의 신호는 Grant Arbiter(60)에서 선택한 input port의 위치 정보를 나타낸다. 이 정보는 +1

Adder(64)에 의해 modular n 형태로 1 증가된 후 Grant Arbiter gj (65)에 전달된다. gj (65)에 전달된 값은 Accept Signal(66)를 받을 경우 gj (65)에 저장된다. Accept Signal(66)은 Grant Arbiter(60)가 생성한 grant 신호가 accept 되었음을 알려주는 신호이다.

<45> Round Robin Selector(62)는 결과신호를 Grant Signal Decoder(68)로 전달한다. Grant Signal Decoder(68)는 SDB State Signal(67)에 따라서 n bit 크기의 Grant Signal Vector(69)를 생성한다. SDB State Signal(67)은 [도 2]의 SDB(24)의 상태를 나타내는 신호로서 SDB가 셀을 받아들일 수 있을 경우 0, SDB가 셀을 받아들일 수 없을 경우 1로 표시된다. Grant Signal Decoder(68)는 SDB State Signal(67)이 1일 경우 Grant Signal Vector(69)를 모두 0으로 set하며, SDB State Signal(67)이 0일 경우 Round Robin Selector(62)에서 선택한 bit 만을 1로 set하고 나머지는 0으로 set하여 출력한다. 이는 Grant Arbiter(60)는 해당되는 output port의 SDB가 셀을 받아 들일 수 있을 경우에만 우선순위에 따라서 n개의 accept 신호 중 하나를 선택하여 grant 신호를 생성함을 의미한다.

<46> [도 7]은 본 발명의 XSU(13)의 Accept Arbiter(70)에 대한 장치구현 실시 예를 나타낸 것이다. Grant Signal Vector(71)가 XSU(13)의 Accept Arbiter(70)에 전달된다. Grant Signal Vector(71)는 n bit 크기로 XSU(13)의 output port에서 특정 input port로 전달하는 grant 신호를 의미한다.

<47> Grant Signal Vector(71)는 Round Robin Selector(72)에 연결된다. Round Robin Selector(72)는 Accept Arbiter의 ai (75)에 저장되어 있는 highest

priority value로부터 round-robin 형태로 Grant Signal Vector(71)를 검색하여 처음으로 grant된 신호 하나를 선택한다. Round Robin Selector(72)는 선택된 grant 신호에 대한 결과로서 $\log(n)$ 크기의 신호를 출력한다. $\log(n)$ 크기의 신호는 Accept Arbiter(70)에서 선택한 output port의 위치 정보를 나타낸다. 이 정보는 +1 Adder(74)에 의해 modular n 형태로 1 증가된 후 ai (75)에 저장된다.

<48> Round Robin Selector(72)는 결과신호를 Accept Signal Decoder(76)로 전달한다. Accept Signal Decoder(76)는 $\log(n)$ 크기의 결과신호에 해당되는 bit 만을 1로 set하고 나머지는 0으로 set하여 n bit 크기의 Accept Signal Vector(77)를 생성한다.

<49> [도 8]은 단일 XSU(13)에 적용되는 XSU Arbiter(80)의 장치구현 실시 예를 나타낸 것이다. XSU Arbiter(80)는 n_2 bit의 State of VOQ (81), n 개의 분산된 Grant Arbiter (60-1) ~ (60-n), n 개의 분산된 Accept Arbiter (70-1) ~ (70-n), Crossbar Switch Fabric Controller(82), 그리고 SDB Controller(85)로 구성된다.

<50> n_2 bit의 State of VOQ (81)는 XSU(13)에 접속되는 n 개의 input port 별 가지고 있는 n 개의 VOQ의 현재 상태를 나타낸 것으로 VOQ가 non-empty일 경우 1이며 empty 일 경우 0의 값을 가진다. 각 Grant Arbiter (60-1) ~ (60-n)은 [도 6]의 Request Signal Vector(61)에 해당하는 n bit의 정보를 State of VOQ(81)로부터 전달 받아 [도 6]의 Grant Signal Vector(69)에 해당하는 n bit 신호를 생성하여 Accept Arbiter (70-1) ~ (70-n)로 전달한다.

<51> 각 Accept Arbiter (70-1) ~ (70-n)은 [도 7]의 Grant Signal Vector(71)에

해당하는 n bit 의 정보를 전달 받아 [도 7]의 Accept Signal Vector(77)에 해당하는 n bit 신호를 생성하여 Crossbar Switch Fabric Controller(82)로 전달한다.

<52> Crossbar Switch Fabric Controller(82)는 $n \times n$ crossbar switch fabric의 각 cross-point를 제어하기 위한 n^2 bit의 crossbar control signals(84)과 output port 별 accept 상태를 파악하여 n bit의 Accept Signal(83)를 생성한다. Crossbar Switch Fabric Controller(82)는 Accept Arbiter (70-1) ~ (70- n)로부터 전달되는 accept signal로부터 output port 별 accept 상태를 파악하여 Accept Signal(83)의 해당되는 bit를 1로 set 한다.

<53> SDB Controller(85)는 SDB의 상태를 파악하여 SDB의 상태가 셀을 저장할 수 있을 경우 0, 셀을 저장할 수 없을 경우 1인 SDB State Signal(86)을 생성하는 기능을 제공한다. SDB Controller(85)는 SDB내에서 1셀 이상의 셀이 있을 경우 Request Signal Vector to Output Arbiters(87) 신호를 생성하여 Output Arbiter에 전달한다. Accept Signal Vector from Output Arbiters(87)신호는 Request Signal Vector to Output Arbiters(87) 신호에 대한 Output Arbiter의 결과 신호로서 이를 이용하여 SDB Controller(85)는 다음 time slot에서 1 cell이 빠져나갈 SDB를 구별할 수 있다. SDB Controller(85)는 Accept Signal(83)로부터 다음 time slot에서 1 cell이 저장 될 SDB를 구별할 수 있다. SDB Controller(85)는 Accept Signal(83)과 Accept Signal Vector from Output Arbiters(87)신호를 사용하여 SDB의 상태를 파악할 수 있다.

<54> [도 9]는 본 발명의 output port에 적용 가능한 Output Arbiter (90)의 장치

구현 실시 예를 나타낸 것이다. Request Signal Vector to Output Arbiters(91)가 Output Arbiter(90)에 전달된다. Request Signal Vector to Output Arbiters(91)는 l bit 크기로 l 개의 XSU에서 Output Arbiter(90)에 전달하는 request 신호이다. Request Signal Vector to Output Arbiters(91)는 Round Robin Selector(92)에 연결된다. Round Robin Selector(92)는 oj (95)에 저장되어 있는 highest priority value로부터 round-robin 형태로 Request Signal Vector to Output Arbiters(91)를 검색하여 처음으로 request된 신호 하나를 선택한다. Round Robin Selector(92)는 선택된 request 신호에 대한 결과로서 $\log(l)$ 크기의 신호를 출력한다. $\log(l)$ 크기의 신호는 Output Arbiter (90)에서 선택한 XSU의 위치 정보를 나타낸다. 이 정보는 +1 Adder(94)에 의해 modular l 형태로 1 증가된 후 ai (75)에 저장된다.

<55> Request Signal Vector to Output Arbiters(91)는 결과신호를 Accept Signal Decoder(96)로 전달한다. Accept Signal Decoder(96)는 $\log(l)$ 크기의 결과신호에 해당되는 bit 만을 1로 set하고 나머지는 0으로 set하여 l bit 크기의 Accept Signal Vector from Output Arbiters(97)를 생성한다.

<56> [도 10]은 본 발명에 적용 가능한 $N \times N$ Switch에 대한 Arbiter(100)의 실시 예를 나타낸 것이다. $N \times N$ Switch에 대한 Arbiter(100)는 L2 개의 분산된 XSU Arbiter(80-1) ~ (80- l 2)와 N 개의 분산된 Output Arbiter(90-1) ~ (90-N)를 interconnection 함으로 구성된다. 각 XSU Arbiter(80-1) ~ (80- l 2)에서 출력되는 Request Signals to Output Arbiters (102-1) ~ (102- l 2)는 Output Arbiter(90-1) ~ (90-N)에 전달되며, 각 Output Arbiter(90-1) ~ (90-N)에서 출력

되는 Accept Signals from Output Arbiters(104-1) ~ (104-N)는 XSU Arbiter(80-1) ~ (80-12)에 전달이 된다.

<57> 비록, 본 발명이 가장 실제적이며 바람직한 실시예를 참조하여 설명되었지만, 본 발명은 상기 개시된 실시예에 한정되지 않으며, 후술되는 특허청 구범위 내에 속하는 다양한 변형 및 등가물들도 포함한다.

【발명의 효과】

<58> N x N Switch에 대한 arbitration은 L2 개의 분산된 XSU Arbiter와 N 개의 분산된 Output Arbiter를 interconnection 함으로 구성되며, 단일 XSU Arbiter는 n 개의 분산된 Grant Arbiter 와 n 개의 분산된 Accept Arbiter로서 구성 가능하므로 본 발명의 경우 scalability 가 매우 높다. 또한 모든 arbitration이 고속구현에 가장 적합한 방식인 round-robin 기반으로 실현이 가능하다. 따라서 본 발명은 소규모 Crossbar Switch Unit들을 상호 연결하여 대규모 스위치를 구성할 수 있는 방식과 이러한 대규모 스위치에 적합한 분산 arbitration 방식에 관한 것으로 고속 대용량의 스위치를 구성하기 위한 필수 조건인 높은 scalability와 분산된 arbitration 을 제공할 수 있다.

<59> 본 방식에 대해서 single iteration in a time slot 만을 가지고 uniform traffic에 대한 시뮬레이션을 수행해보면 동일한 조건의 타 방식(PIM, iSLIP)에 비해서 상대적으로 우수한 셀 전달 지연시간 성능을 가지면서 100%의 throughput을 전달할 수 있음을 확인할 수 있으므로 타 방식에 비해서 상대적으로 우수한

1020020083720

arbitration time margin을 가지고 있음을 알 수 있다.

【특허청구범위】

【청구항 1】

N개의 Input Port와 N개의 Output Port와 $N \times N$ switch fabric으로 구성되는 스위치에서, 각 Input Port는 N개의 Virtual Output Queue (이후 VOQ로 표기)를 가지고, n은 N의 약수이며 $L=N / n$ 의 관계가 있을 때 N 개의 VOQ는 순차적으로 n개의 VOQ 단위로 묶어서 물리적으로 분리된 L개의 VOQ 그룹을 구성하며, L개의 VOQ 그룹은 각각 독립된 switch fabric interface port를 통하여 물리적으로 분리된 L개의 XSU에 각각 연결되고, $N \times N$ 크기의 switch fabric은 L2 개의 crossbar switch unit (이후 XSU로 표기)으로 구성되며, 각 XSU는 독립된 arbitration 기능을 통하여 n개의 input ports로부터 들어오는 셀을 n개의 Output Port로 전달하는 기능을 수행하고, 각 Output Port는 L개의 switch fabric interface port를 통하여 물리적으로 분리된 L개의 XSU에 연결되어 있으며 서로 독립된 arbitration을 통하여 a cell time slot 동안 L개의 XSU로부터 1개의 셀을 선택하여 출력 링크로 전달하는 기능을 제공하는 것을 특징으로 하는 장치.

【청구항 2】

제 1 항에 있어서,

XSU은 n개의 input ports와 n개의 output ports를 가지는 $n \times n$ crossbar switch fabric과, 각 output port에 일정크기의 셀을 저장할 수 있는 small size scheduling decomposition buffer(이후 SDB로 표기)로 구성되어 있고, XSU의 각

output port는 SDB 가 셀을 받아 들일 수가 있을 경우 XSU output port별 독립된 arbitration을 통하여 해당 SDB에 셀을 저장하는 장치.

【청구항 3】

제 1 항에 있어서,

하나의 crossbar switch module (이후 XSM으로 표기) 은 L개의 XSU으로 구성되며, 하나의 XSM은 n개의 Input Port로부터 셀을 받아서 N개의 Output Port로 전달하는 기능을 제공함에 있어서 m 번째 XSM에 연결되는 input port는 ($m \times n$) 번째 input port로부터 $\{m \times (n+1) 1\}$ 번째 input port까지 이고, 단일 XSM과 n개의 input ports는 물리적으로 분리된 Subsystem으로 구성할 수 있고, 이러한 L개의 Subsystem을 이용하여 전체 $N \times N$ 시스템을 구성할 수 있는 modular한 특성을 가지고 있는 장치.

【청구항 4】

제 1 항에 있어서,

하나의 Output Port는 memory-less형태로 구성되며, L개의 SDB로부터 a time slot동안 최대 1개의 셀을 Output Arbiter가 선택하여 output port로 전달하는 기능을 수행함에 있어서 j 번째 출력포트에 해당되는 SDB는 $\sum_{m=0}^{L-1} SDB(m, \left\lfloor \frac{j}{n} \right\rfloor, j \bmod n)$ 인 관계를 특징으로 하는 장치.

【청구항 5】

제 2 항에 있어서,

단일 XSU의 arbitration은 n_2 bit의 State of VOQ, n 개의 분산된 Grant Arbiter, n 개의 분산된 Accept Arbiter, Crossbar Switch Fabric Controller, 그리고 SDB Controller로 구성되고, n_2 bit의 State of VOQ는 XSU에 접속되는 n 개의 input port 별 가지고 있는 n 개의 VOQ의 현재 상태를 나타낸 것으로서, 단일 Grant Arbiter는 n_2 bit의 State of VOQ 중 n bit Request Signal Vector를 전달 받아 n bit Grant Signal Vector 생성하여 n 개의 Accept Arbiter로 전달하고, 단일 Accept Arbiter는 n bit Grant Signal Vector 정보를 전달 받아 n bit Accept Signal Vector 신호를 생성하여 Crossbar Switch Fabric Controller로 전달하고 Crossbar Switch Fabric Controller는 $n \times n$ crossbar switch fabric의 각 cross-point를 제어하기 위한 n_2 bit의 crossbar control signals과 output port 별 accept 상태를 파악하여 n bit의 Accept Signal을 생성하고, SDB Controller는 SDB의 상태를 파악하여 SDB내에 1셀 이상의 셀이 있을 경우 Request Signal Vector to Output Arbiters 신호를 Output Arbiter에 전달하고 그 결과로서 Output Arbiter로부터 전달되는 Accept Signal Vector from Output Arbiters 신호와 Crossbar Switch Fabric Controller로부터 전달되는 Accept Signal을 이용하여 SDB의 상태를 파악할 수 있는 것을 특징으로 하는 장치.

【청구항 6】

제 5 항에 있어서,

단일 Grant Arbiter는 n개의 input port로부터 전해온 request 신호 중 하나를 선택함에 있어서 Grant Arbiter의 내부 레지스터에 저장되어 있는 highest priority value로부터 round-robin 형태로 검색하여 처음으로 request한 input port를 선택한 후 Grant Arbiter가 해당되는 output port의 SDB가 추가 셀을 받아들일 수 있을 경우에만 해당되는 input port로 grant 신호를 전달하고 그렇지 않을 경우 grant 신호를 전달하지 않으며, 선택된 input port의 port 정보는 modular n 형태로 1 증가된 후 Grant Arbiter 가 grant한 input port로부터 accept 신호를 받을 경우에만 새로운 highest priority value로서 내부 레지스터에 저장하고 그렇지 않을 경우 이전 highest priority value를 그대로 유지하는 것을 특징으로 하는 장치.

【청구항 7】

제 5 항에 있어서,

단일 Accept Arbiter는 n개의 output port로부터 전해온 grant 신호 중 하나를 선택함에 있어서 Accept Arbiter의 내부 레지스터에 저장되어 있는 highest priority value로부터 round-robin 형태로 검색하여 처음으로 grant한 output port를 선택한 후 해당되는 output port로 accept 신호를 전달하고, 선택된 output port의 port 정보를 modular n 형태로 1 증가한 후 새로운 highest priority value

로서 내부 레지스터에 저장하는 것을 특징으로 하는 장치.

【청구항 8】

제 1 항에 있어서,

단일 Output Arbiter는 1개의 XSU으로부터 전해온 request 신호 중 하나를 선택함에 있어서 Output Arbiter의 내부 레지스터에 저장되어 있는 highest priority value로부터 round-robin 형태로 검색하여 처음으로 request한 XSU를 선택한 후 해당되는 XSU로 accept 신호를 전달하고, 선택된 XSU의 정보를 modular n 형태로 1 증가한 후 새로운 highest priority value로서 내부 레지스터에 저장하는 것을 특징으로 하는 장치.

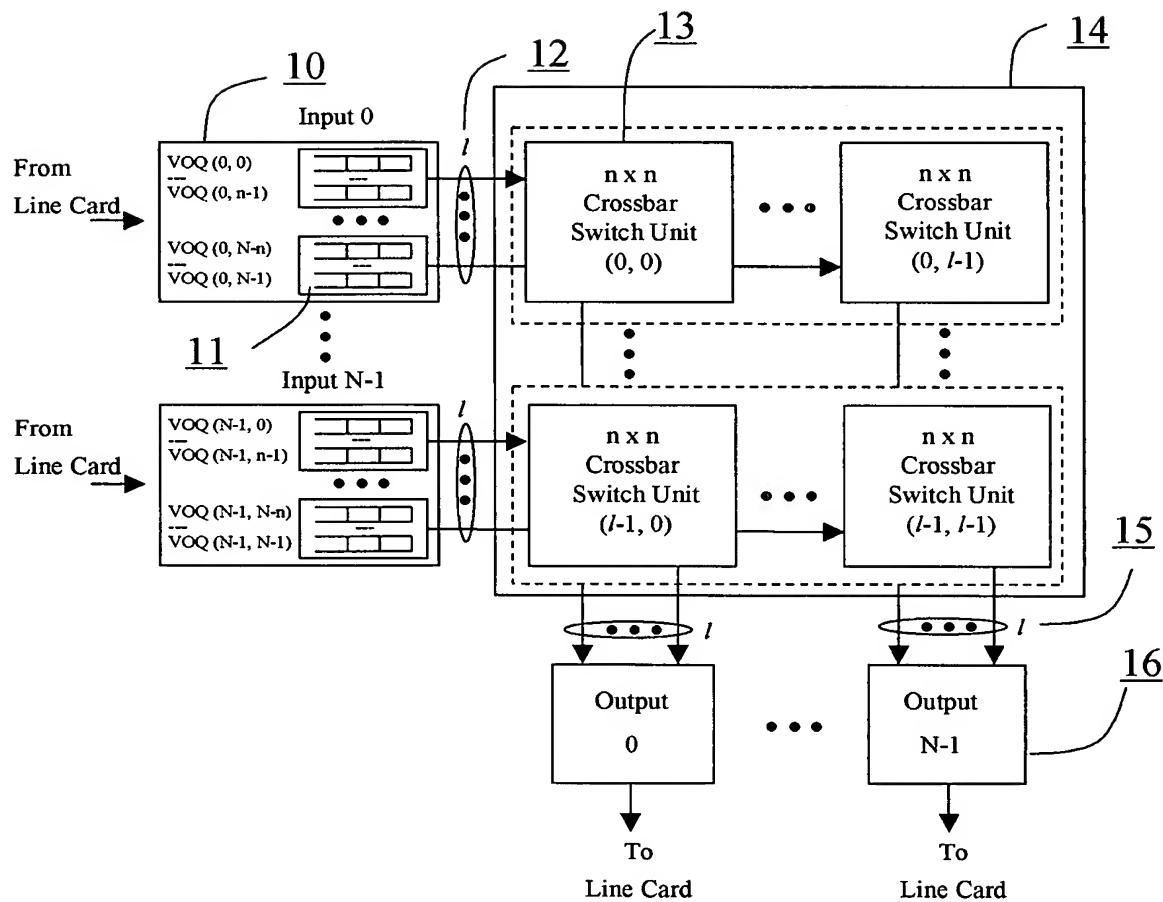
【청구항 9】

제 1 항에 있어서,

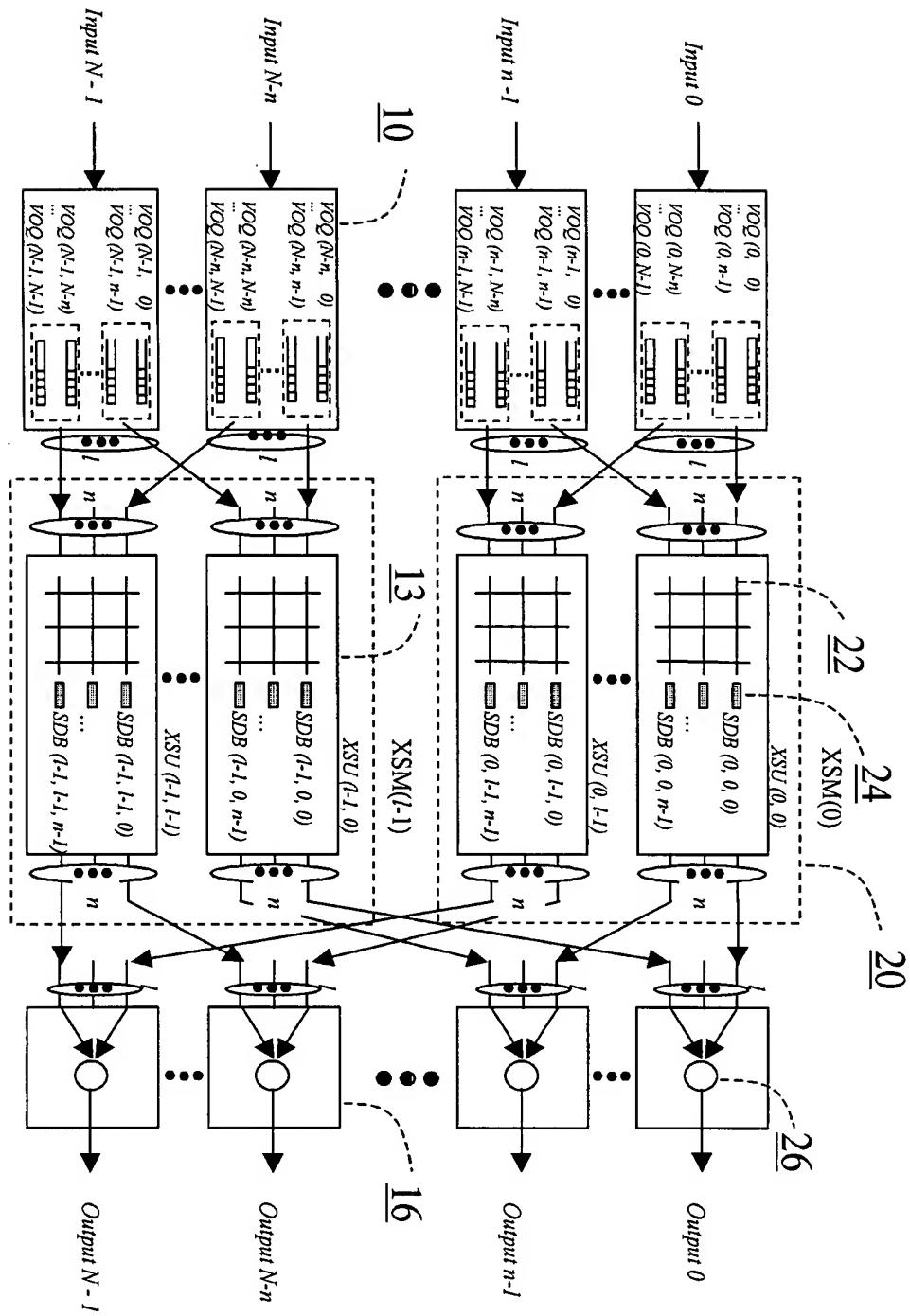
$N \times N$ Switch에 대한 Arbiter를 L2 개의 분산된 XSU Arbiter와 N 개의 분산된 Output Arbiter를 interconnection 함으로 구성함에 있어서 L2 개의 분산된 XSU는 Request 신호를 생성하여 Output Arbiter에 전달하며, N개의 Output Arbiter는 Request 신호를 처리한 결과인 Accept 신호를 XSU Arbiter에 전달함에 있어서 단일 XSU Arbiter는 n 개의 분산된 Grant Arbiter 와 n 개의 분산된 Accept Arbiter로서 구성되는 것을 특징으로 하는 장치.

【도면】

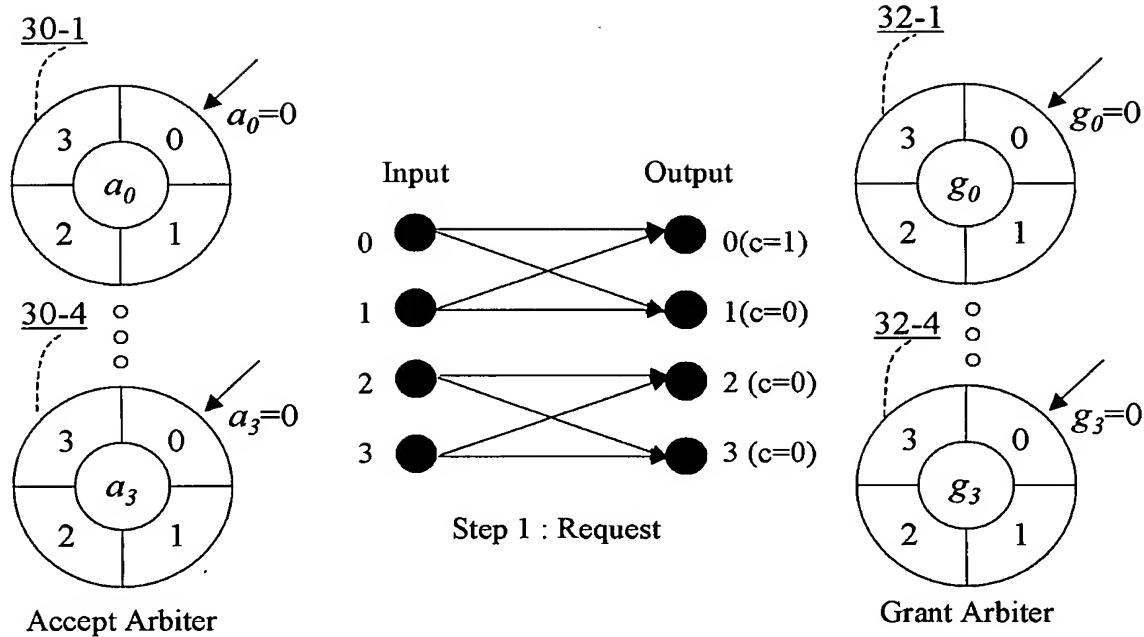
【도 1】



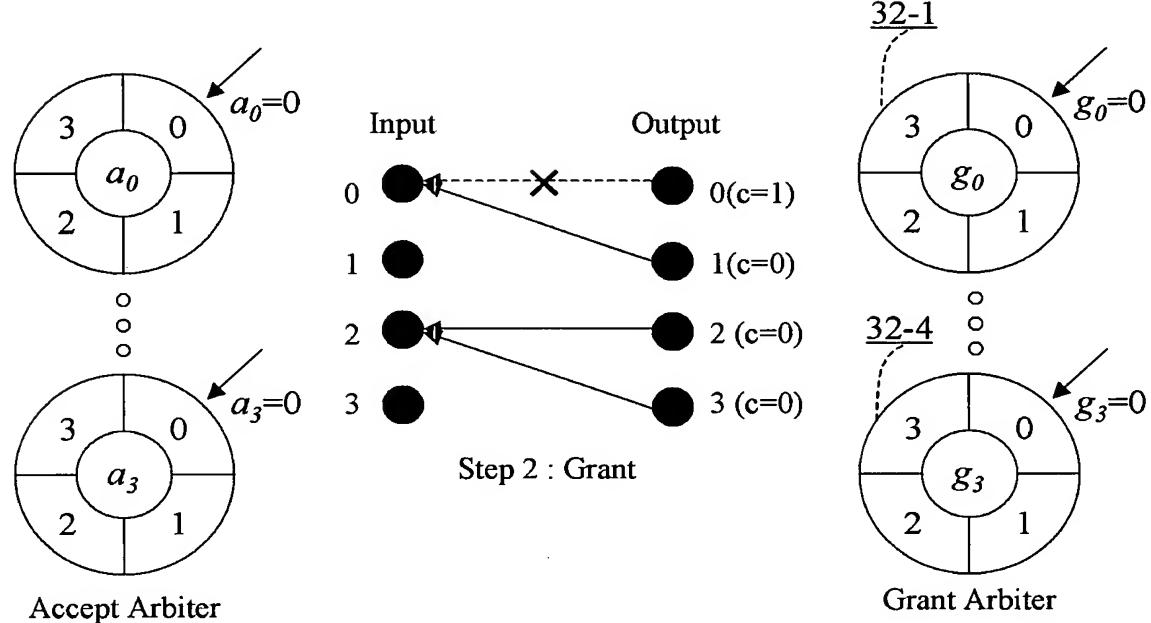
【도 2】



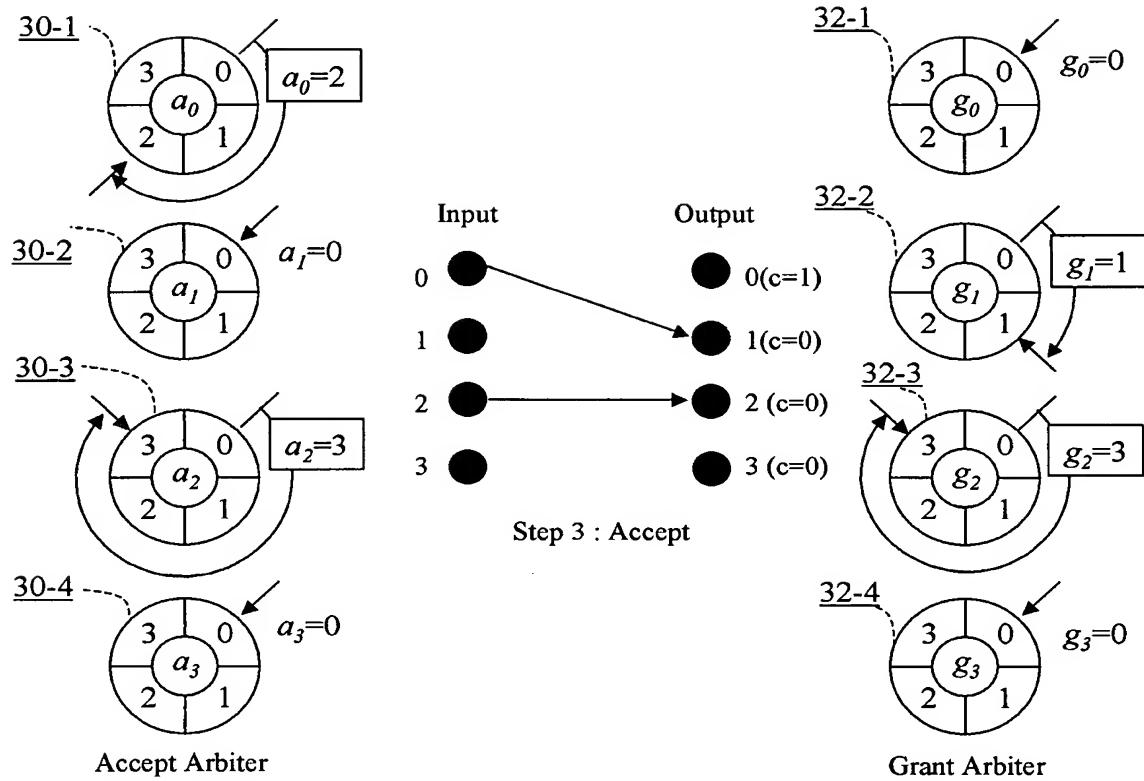
【도 3】



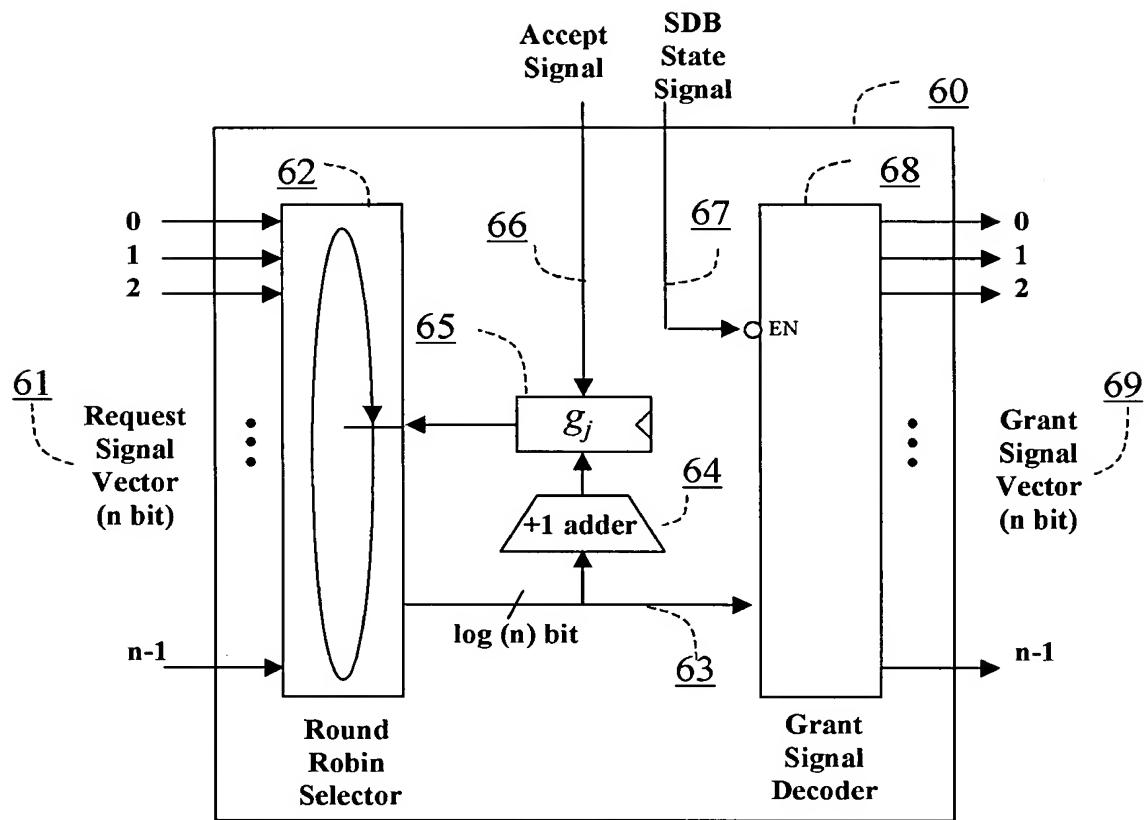
【도 4】



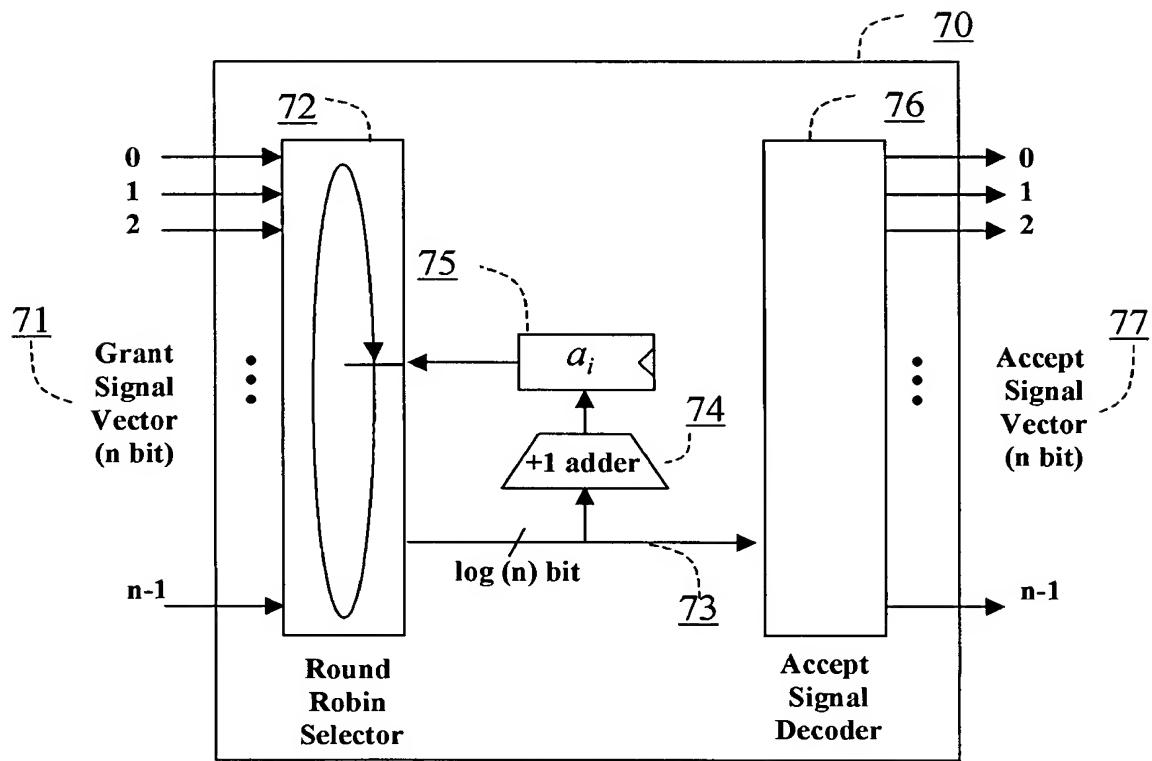
【도 5】



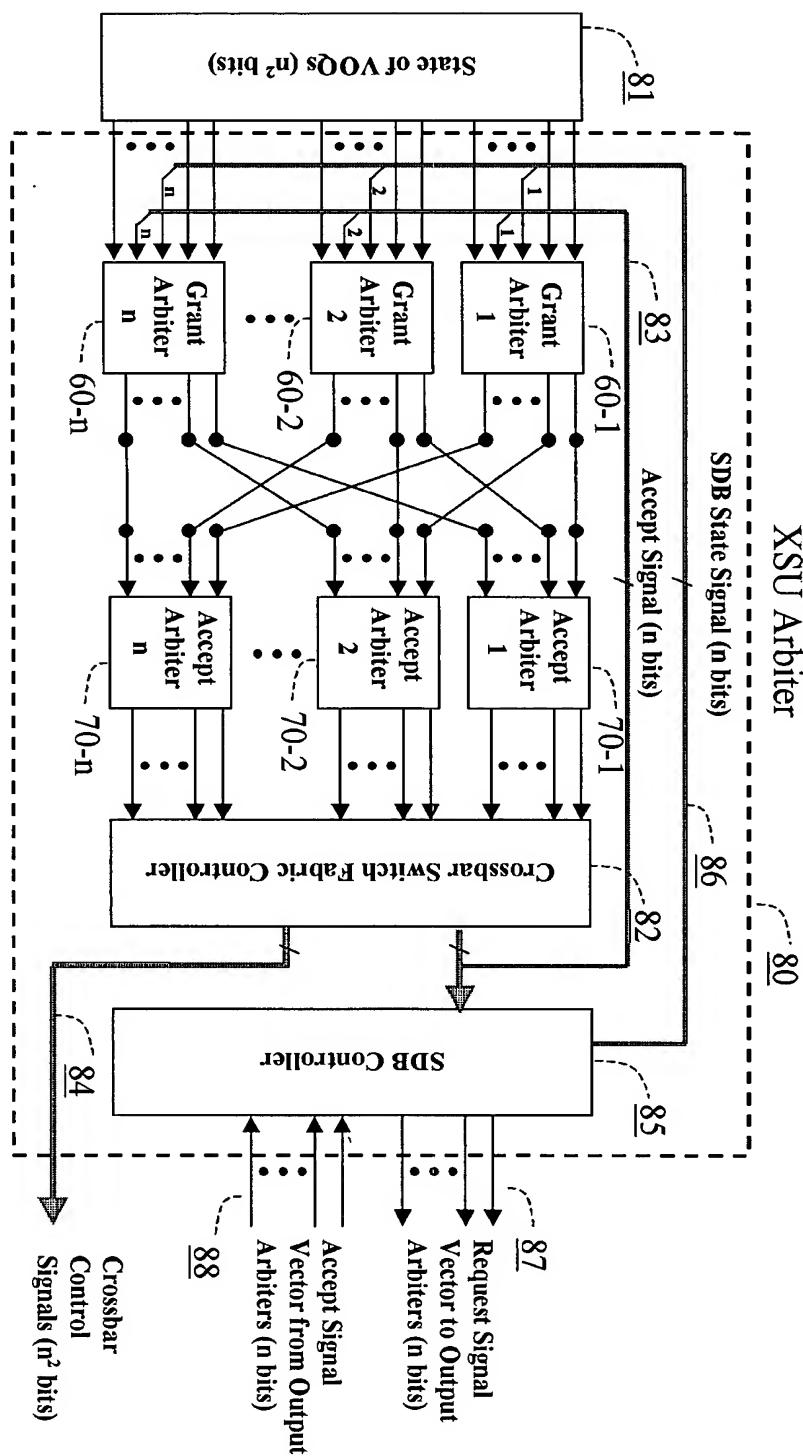
【도 6】



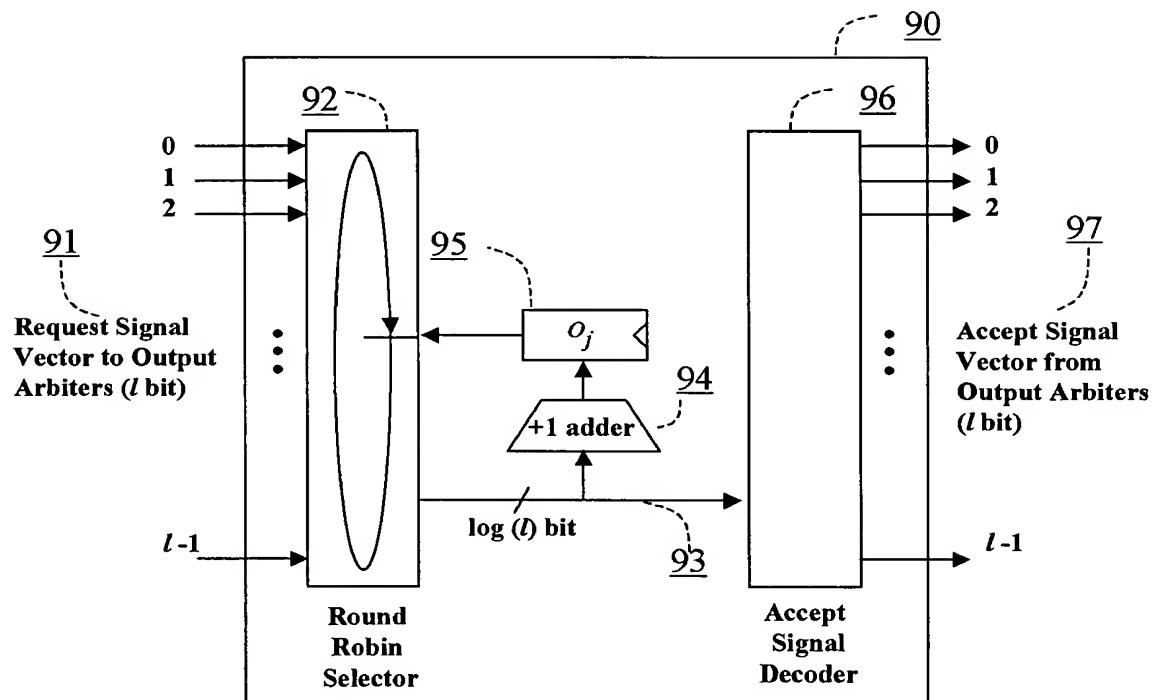
【도 7】



【8-5】



【도 9】



【도 10】

